Connect. Accelerate. Outperform.™

CloudX™ VMware vCloud 5.5

Reference Guide

Rev 1.0

www.mellanox.com

NOTE:
THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT ("PRODUCT(S)") AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES "AS-IS" WITH ALL FAULTS OF ANY KIND AND SOLELY FOR THE PURPOSE OF AIDING THE CUSTOMER IN TESTING APPLICATIONS THAT USE THE PRODUCTS IN DESIGNATED SOLUTIONS. THE CUSTOMER'S MANUFACTURING TEST ENVIRONMENT HAS NOT MET THE STANDARDS SET BY MELLANOX TECHNOLOGIES TO FULLY QUALIFY THE PRODUCT(S) AND/OR THE SYSTEM USING IT. THEREFORE, MELLANOX TECHNOLOGIES CANNOT AND DOES NOT GUARANTEE OR WARRANT THAT THE PRODUCTS WILL OPERATE WITH THE HIGHEST QUALITY. ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Mellanox Technologies
350 Oakmead Parkway Suite 100
Sunnyvale, CA 94085
U.S.A.
www.mellanox.com
Tel: (408) 970-3400
Fax: (408) 970-3403

Mellanox Technologies, Ltd.
Hakidma 26
Ofer Industrial Park
Yokneam 2069200
Israel
www.mellanox.com
Tel: +972 (0)74 723 7200
Fax: +972 (0)4 959 3245

© Copyright 2015. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, ConnectX®, Connect-IB®, CoolBox®, CORE-Direct®, GPUDirect®, InfiniBridge®, InfiniHost®, InfiniScale®, Kotura®, Kotura logo, MetroX®, MLNX-OS®, PhyX®, ScalableHPC®, SwitchX®, TestX®, UFM®, Virtual Protocol Interconnect®, Voltaire® and Voltaire logo are registered trademarks of Mellanox Technologies, Ltd.

ExtendX™, FabricIT™, FPGADirect™, HPC-X™, Mellanox Care™, Mellanox CloudX™, Mellanox Open Ethernet™, Mellanox PeerDirect™, Mellanox Virtual Modular Switch™, MetroDX™, NVMeDirect™, Switch-IB™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Table of Contents

About this Manual **5**

1 Overview **7**

1.1 Solution Components 7

1.1.1 Hardware Requirements 7

1.1.2 Software Components 7

1.1.3 Deployment Components 7

1.2 Setup Diagram 8

2 Network Installation **9**

2.1 Networks Allocation 9

2.2 Configuring 1Gbps WAN/Public Switch 9

2.3 Configuring 40Gbps Switching Fabric (MLAG, VLANs) 9

3 Host Installation and Configuration **11**

4 vShield Manager Appliance Installation **19**

5 VXLAN Configuration **22**

6 Virtual Tunnel Endpoint IP Configuration **24**

7 vCloud Director Deployment **26**

List of Tables

Table 1: Related Documents 5
Table 2: Network Allocation Example 9

About this Manual

Mellanox CloudX™ is a group of reference architectures that allows companies to build the most efficient, high performance and scalable clouds based on Mellanox's superior interconnect and off-the-shelf building blocks (servers, storage, interconnect, and software). This document covers CloudX™ reference architecture for the most efficient cloud infrastructure based on VMware® vCloud solutions while running on Mellanox interconnect technology.

Audience

This document is intended for cloud, virtualization and network architects who are interested in deploying VMware based cloud solutions using Mellanox up to 56GbE NIC and switch products.

Document Conventions

The following lists conventions used in this document.

NOTE: Identifies important information that contains helpful suggestions.

CAUTION: Alerts you to the risk of personal injury, system damage, or loss of data.

WARNING: Warns you that failure to take or avoid a specific action might result in personal injury or a malfunction of the hardware or software. Be aware of the hazards involved with electrical circuitry and be familiar with standard practices for preventing accidents before you work on any equipment.

Related Documents

For additional information, see the following documents:

Table 1: Related Documents

Document	Location
VMware® VXLAN Deployment Guide	http://www.vmware.com/files/pdf/techpaper/VMware-VXLAN-Deployment-Guide.pdf
VMware® Network Virtualization Design Guide	http://www.vmware.com/files/pdf/techpaper/Virtual-Network-Design-Guide.pdf

Document	Location
How To Configure MLAG on Mellanox Switches	https://community.mellanox.com/docs/DOC-1434
MLNX-OS User Manual	http://support.mellanox.com/ NOTE: Active support account required to access manual.

1 Overview

CloudX™ for VMware vCloud is a private implementation of cloud reference design based on VMware vCloud Suite and Mellanox advance interconnect NIC and switch products.

This CloudX takes advantage of Virtual eXtensible Local Area Network (VXLAN) offload capabilities of Mellanox ConnectX®-3 Pro NIC device to provide both the scalability and efficiency of VXLAN without the associated penalties: significant drop of network throughput and increase of CPU utilization.

In addition, this reference design takes advantage of Mellanox 40Gbps end-to-end interconnect to provide ample IO services for the VMs running on each Hyper-V hypervisor increasing the VM density and the overall performance of the cloud.

1.1 Solution Components

1.1.1 Hardware Requirements

- All server platform shall be compatible with the VMware ESXi 5.5 U2 requirements.
- An available PCI Express v3.0 x8
- Mellanox ConnectX-3Pro EN / VPI dual port adapter[1]
- Mellanox SwitchX products with 40Gbps Ethernet support

1.1.2 Software Components

The following components: installation and configuration are covered in this document

- VMware ESXi 5.5 U2
- VMware vCloud Networking and Security 5.5.4-2504419
- VMware vCloud Director 5.5.3-2578696

Prerequisites: We assume the components below are already installed and configured on your system. We will run there components as VMs on dedicated, standalone ESXi server (MISC-MGMT server)

- Active Directory domain controller with DHCP and DNS server (based on Windows Server 2012R2)
- SQL server 2014 (for vCenter and vCloud DB)
- VMware vCenter 5.5 U2

1.1.3 Deployment Components

- Switch Fabric:
  - Six MSX6710 switches for cloud internal networks[2]
  - 1Gbps switch for WAN/External network connectivity
- Two compute clusters Gold and Bronze in the same vCenter datacenter:

[1] For Bronze and Gold clusters a single port adapter can be used too.
[2] Mellanox SwitchX VPI switches can be used too.

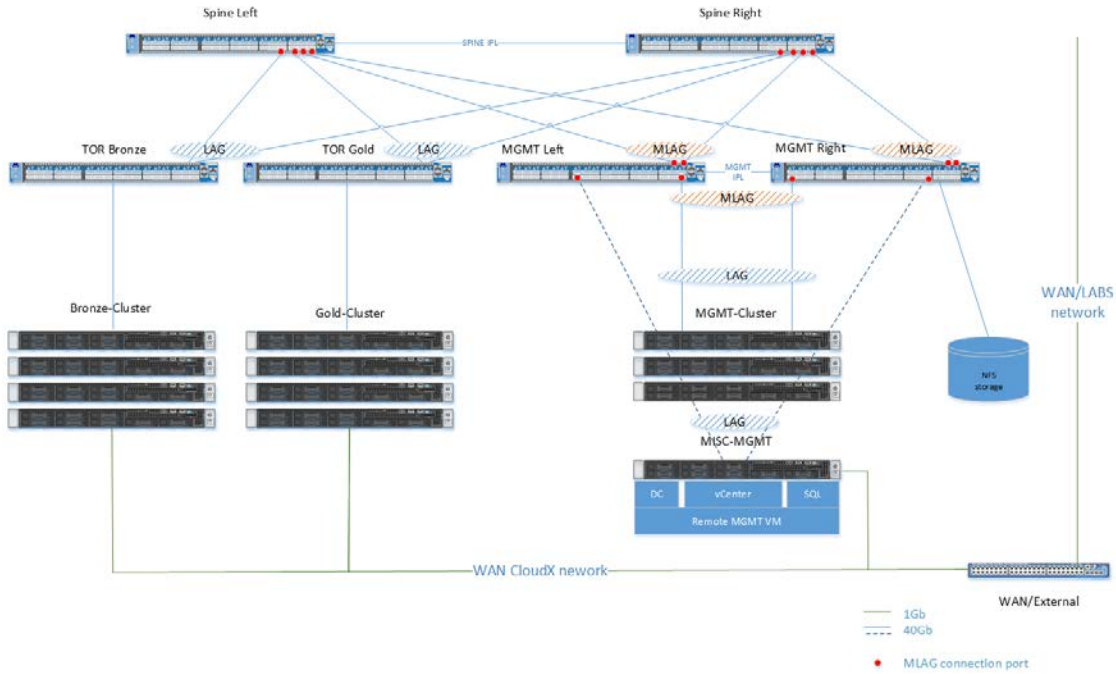
- Four hosts per cluster
- One Top of the Rack switch per cluster
- One 40/56GbE ConnectX®-3 Pro VPI Dual-Port Adapter network interface cards per host[1]
- Two Virtual Distributed Switches (External, VXLAN)

- Management cluster:
  - 3 hosts in the cluster
  - 2 Top of the Rack switches
  - vCloud Director
  - vCloud Networking and Security Manager
  - One Virtual Distributed Switch

1.2 Setup Diagram

Figure 1 presents the cloud topology.

Figure 1: Cloud Topology



Mellanox Technologies Confidential

2 Network Installation

2.1 Networks Allocation

The example in this document is based on the network allocation defined in this table:

Table 2: Network Allocation Example

Network	Switching Fabric	Subnet/Mask	Gateway	VLAN ID
Management and Storage	Mellanox 40Gbps	10.142.200.0/24	10.142.200.1	101
vMotion	Mellanox 40Gbps	10.142.201.0/24	N/A	102
VXLAN	Mellanox 40Gbps	10.142.202.0/24	N/A	103
WAN/Public	1Gbps switch	10.144.255.41-53/16	10.144.0.1	Public network used to connect Cloud nodes to an external network. **Note:** In the examples used in this document, Lab network is used as WAN/Public.

2.2 Configuring 1Gbps WAN/Public Switch

- No specific configuration is required
- Each server in Bronze and Gold clusters should be connected to the network
- The range of IP addresses within the WAN/Public network to be used by the cloud should be allocated by the Network administrator

2.3 Configuring 40Gbps Switching Fabric (MLAG, VLANs)

- **Left and Right spine switches**
  - Set the MTU to at least 1600 for all ports. In our setup we set to 9000.
  - Disable the STP protocol
  - Create an MLAG between the spine switches
  - Set up an `"mlag-port-channel"` interface for each uplink from the TOR switch
  - Set all `"mlag-port-channel"` interfaces to trunk mode with pass-through all VLANs
  - Save the running configuration
- **Bronze and Gold cluster switches**
  - Set the MTU to at least 1600 for all ports. In our setup we set to 9000.
  - Disable the STP protocol
  - Create a LAG interface between the uplinks and the spine switches
  - Save the running configuration

- **Management cluster switches**
  - Set the MTU to at least 1600 for all ports. In our setup we set to 9000.
  - Disable the STP protocol
  - Create an MLAG between the management switches
  - Set up an `"mlag-port-channel"` interface for each uplink from the TOR switch to the spine switch
  - Set all `"mlag-port-channel"` interfaces to trunk mode with pass-through all VLANs
  - Create an `"mlag-port-channel"` interface for each management server connection port and set it to hybrid mode with default VLAN 101 and tagged VLAN 102
  - Save the running configuration

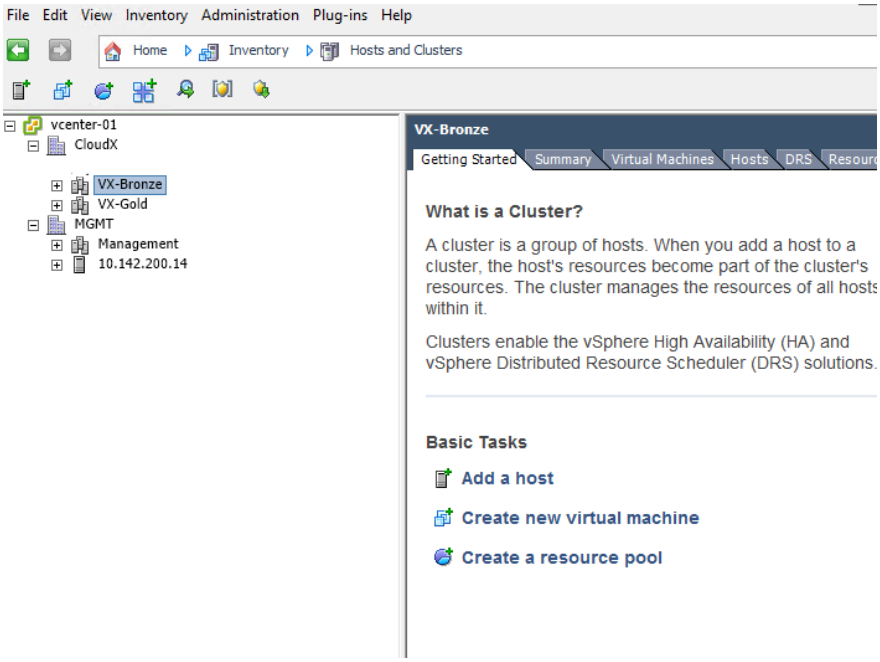
NOTE: For further information of switch configuration, please refer to the MLNX-OS User Manual (See Table 1: Related Documents).

Mellanox Technologies Confidential

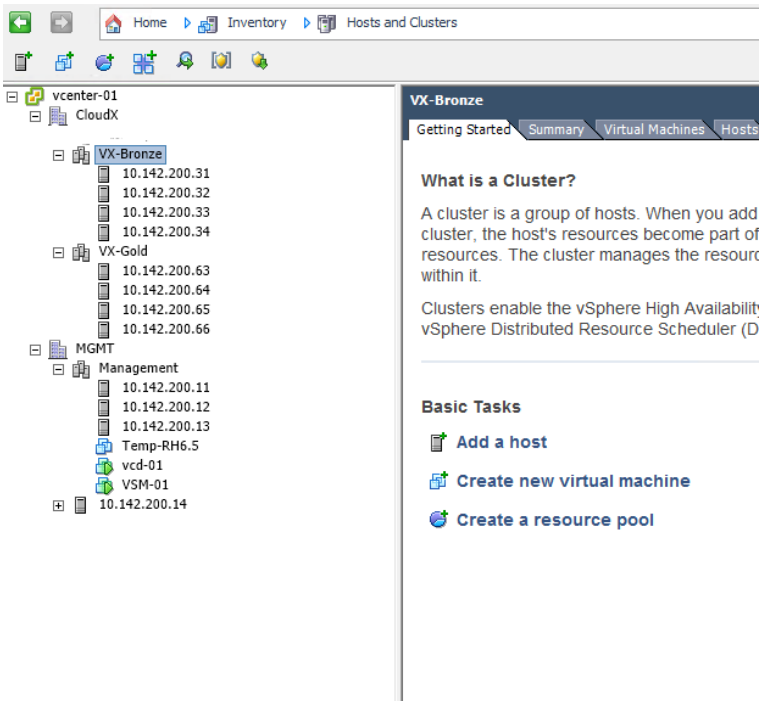
3 Host Installation and Configuration

In order to deploy the host, please follow these steps:

1. Install ESXi 5.5 U2 on all the servers.
2. Configure vCenter server to connect to each ESXi host for GUI/WEB management.
3. Log into your vCenter Server via the vSphere Client or vSphere Web Client and create two datacenters (MGMT, CloudX) with three clusters (Management, VX-Bronze, VX-Gold).



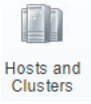
4. Drag your hosts to the corresponding cluster.



5. Manually upgrade the MLNX-EN-ESX Driver for VMware ESX 5.5.

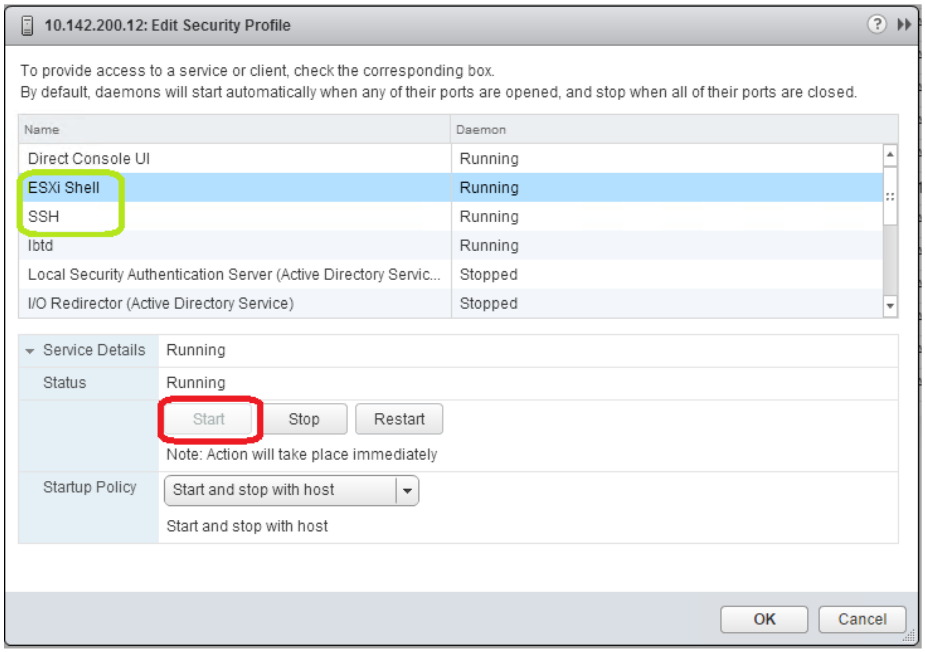
   a. Download the latest Mellanox driver from VMware site. In our example we use MEL-mlnx-en-1.9.9.4-1OEM.550.0.0.1331820-offline_bundle-2349568.zip.

   b. Enable SSH and ESXi Shell on each host in Security Profile services via vSphere Web Client.

   (Select Hosts and Clusters 

→ Select the desired server → Manage Tab → Settings Tab → Security Profile → Scroll down to the Services section and press Edit)

   In edit window start ESXi Shell and SSH services.



   Click **OK**.

   c. Copy the driver to each host in the /tmp folder using a secure shell copy application.

   For example WinSCP .

   d. Install the driver.

   Run from each host's console the following command:

   ```
   "esxcli software vib install -d /tmp/MEL-mlnx-en-1.9.9.4-
   1OEM.550.0.0.1331820-offline_bundle-2349568.zip & reboot"
   ```

6. Create the distributed switches.

   - **NSX_DVSwitch** – A single uplink with two port groups NSX-MGMT, NSX_vMotion
   - **WAN_DVSwitch** – A single uplink with port group WAN_DPG
   - **MGMT_DVS** - Two active uplinks (Load balancing: Route based on IP hash) with two port groups DPortGroup_MGMT, DPortGroup_vMotion

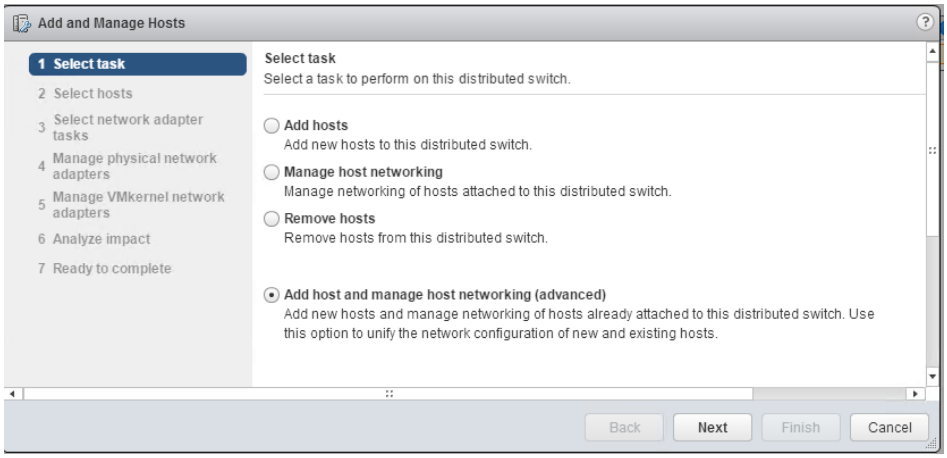
     Note: Networks CloudX-LAN and LAB Network used by the Misc management server and without specific settings.

Mellanox Technologies Confidential

7. Migrate management interface from the standard vSwitch to **NSX_DVSwitch** vSphere distributed switch (DVS).

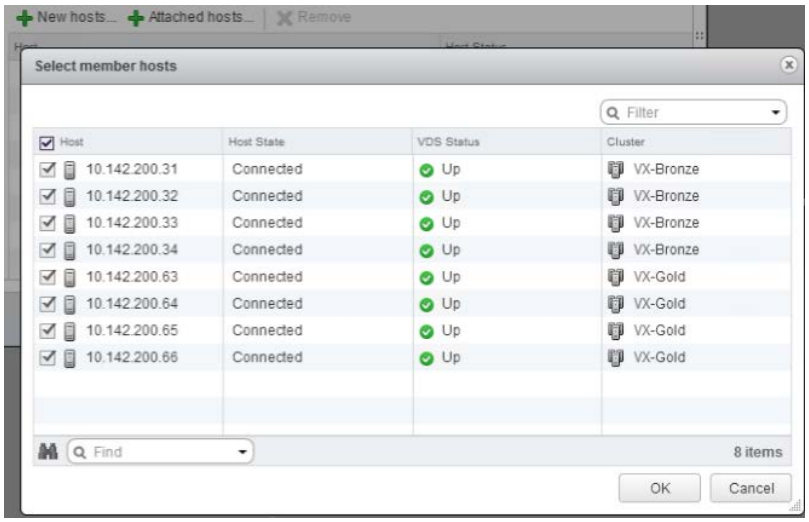
   (Select Networking → Distributed Switch → Manage Tab → Settings Tab → Topology → Add and Manage Hosts)

   a. Choose Add host and manage host networking (advanced)



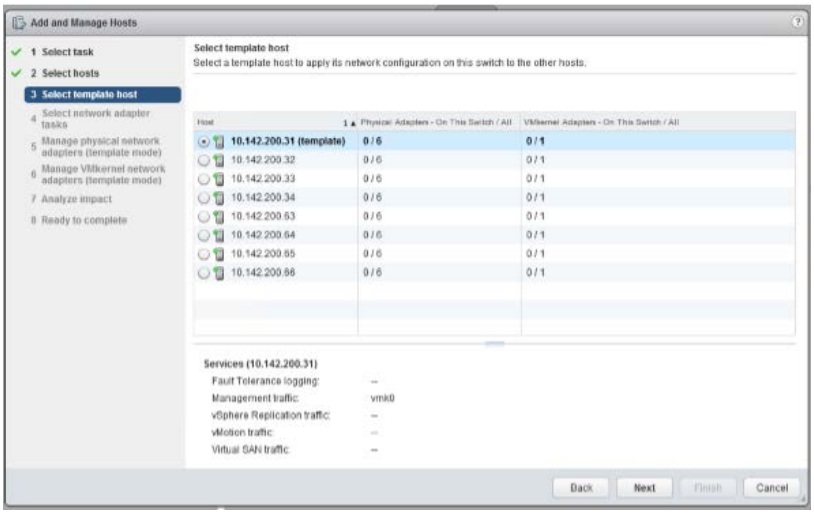
   Click **Next**.

   b. Press **New hosts…** and select hosts.



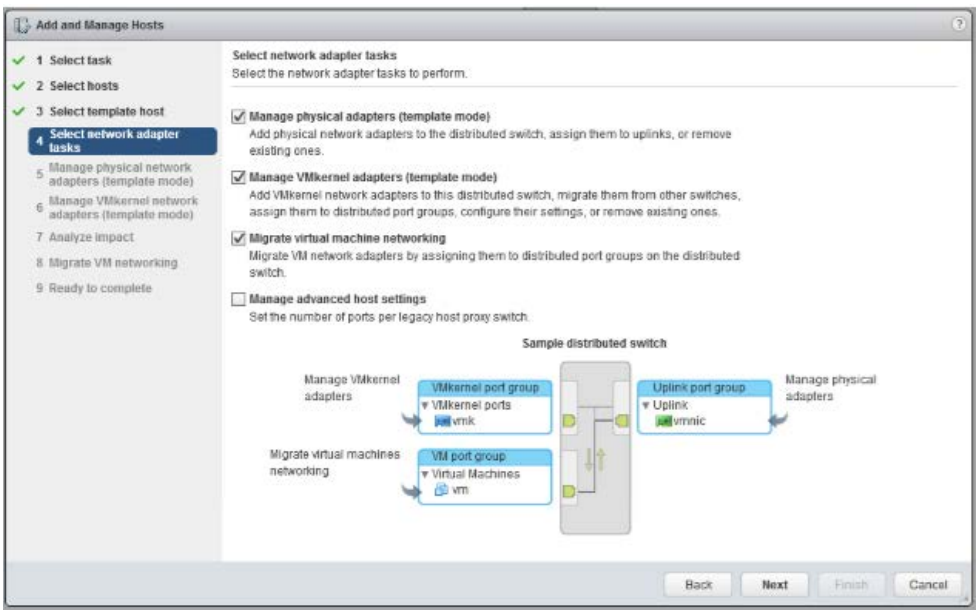
   Click **OK** and **Next**.

c. Select any host as your Template host.



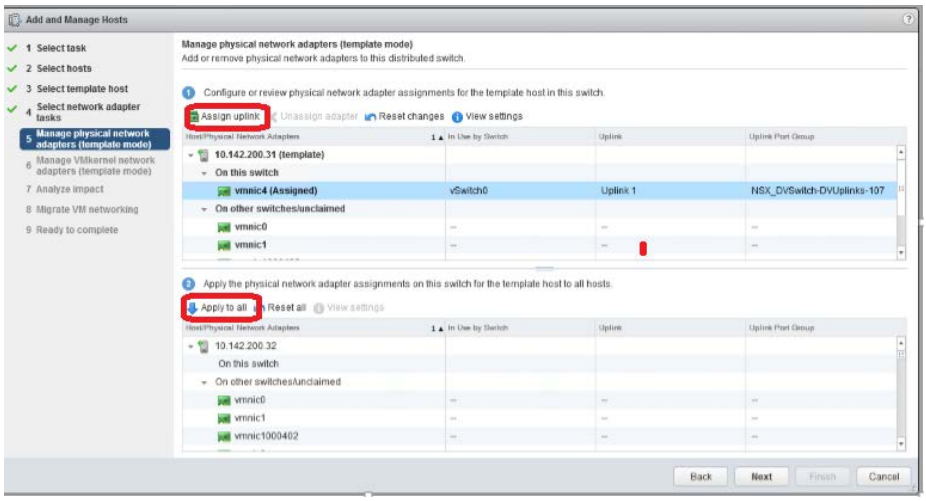
Click **Next**.

d. Choose the first 3 options.



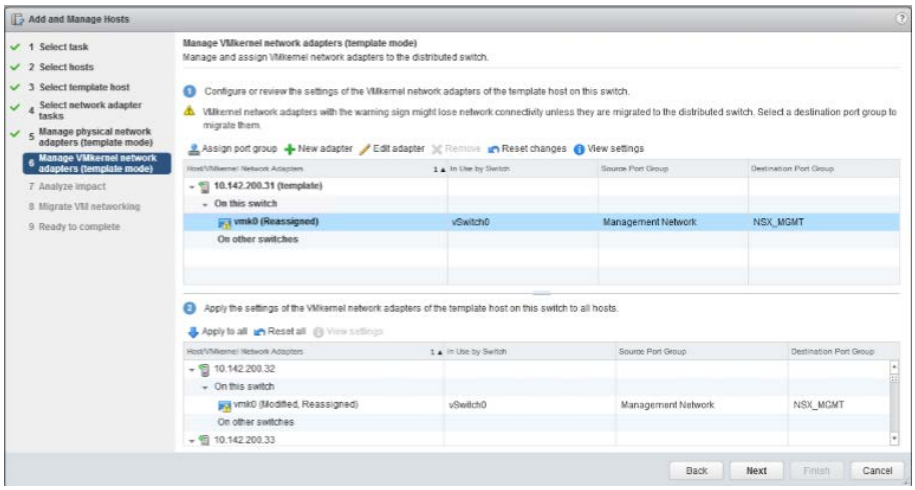
Click **Next**.

e. Assign the Mellanox interface as the uplink on the template host and press **Apply to all.**

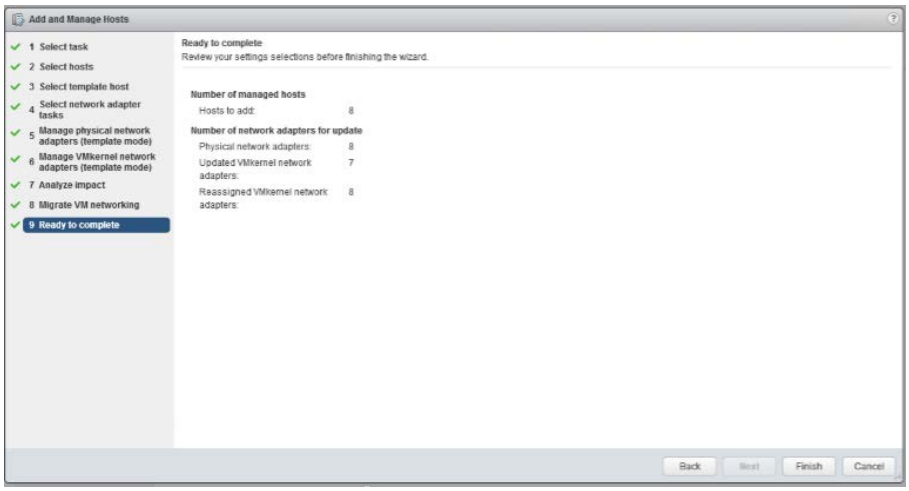


Click **Next**.

f. Assign an NSX_MGMT port group and press **Apply to all.**



Click **Next** until **Finish**.

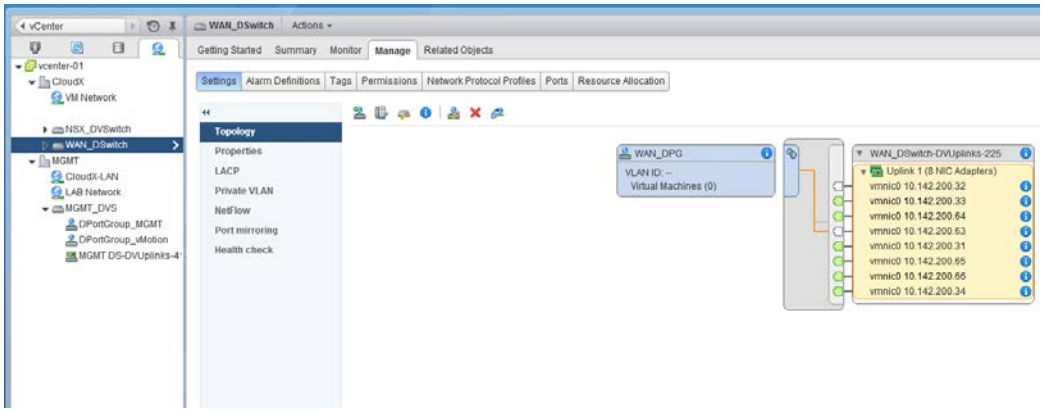


8. Set up the **WAN_DSwitch** vSphere distributed switch.

   Repeat steps 6.a – 6.f with the following changes:

   - In “Step e” choose 1 Gbps interface connected to WAN/Public network.
   - In “Step f” choose WAN_PG port group.

   Once completing these steps, the window below will display your results.

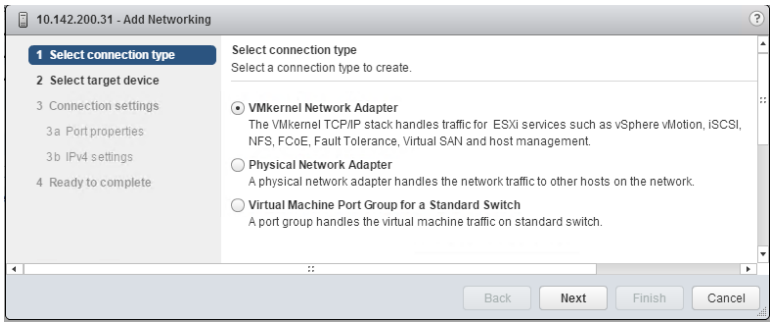


9. Configure VMkernel interface for vMotion Port Group on **each** server in both Bronze and Gold cluster.

   (Click the Home button on vSphere Web Client and select Hosts and Clusters → CloudX datacenter → Related Objects → Hosts Tab)

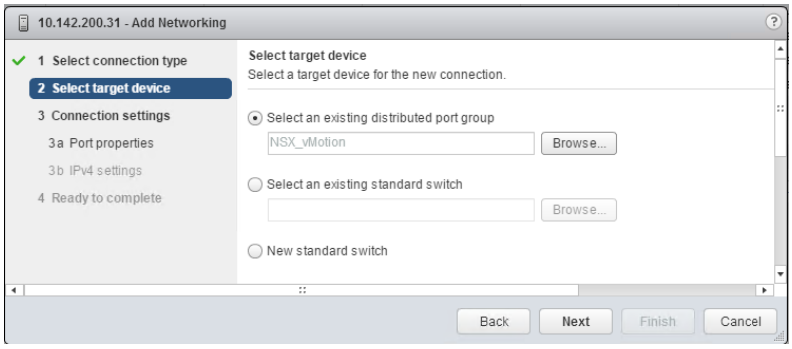
   a. Select the desired host.

   b. Select the Manage Tab → Networking → VMkernel adapters and add the VMkernel Network Adapter .



   Click **Next**.

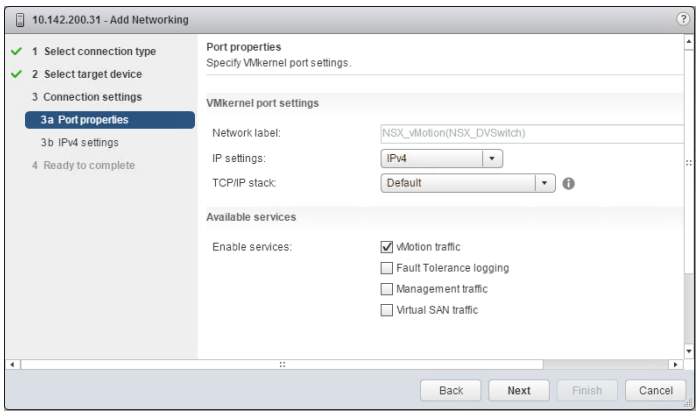
   c. Select target port group.



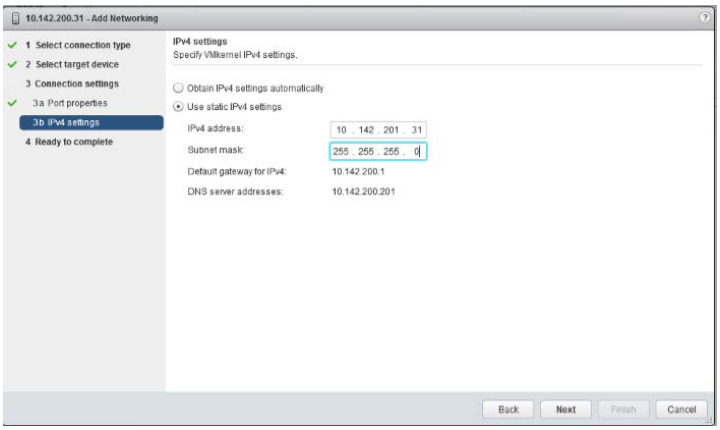
Mellanox Technologies Confidential

Click **Next**.

d. Enable **vMotion traffic** and click **Next**.



e. Specify VMkernel IPv4 settings.



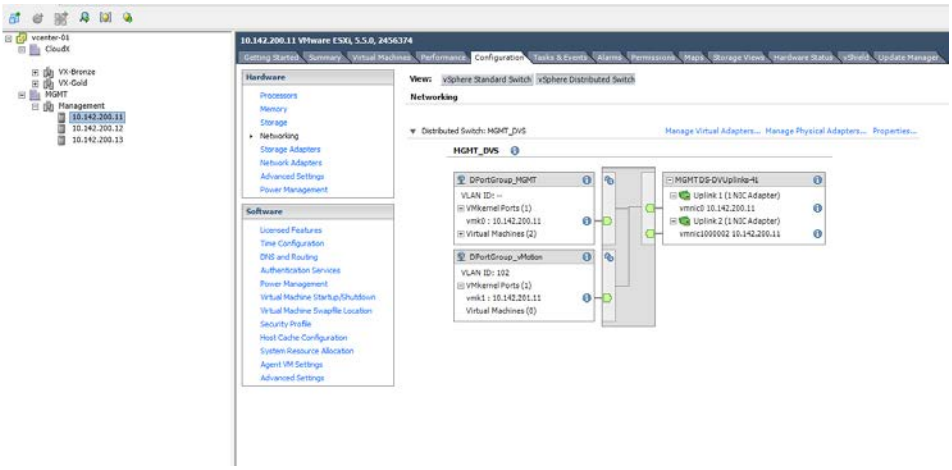
Click **Next** and **Finish**.

10. Migrate management interface from the standard vSwitch to **MGMT_DVS** vSphere distributed switch and set up VMkernel interface for vMotion Port Group in the Management cluster.
    Repeat steps 6.a – 6.f with the following changes:

    - In “Step e” choose two Mellanox interfaces.
    - In “Step f” choose DPortGroup_MGMT port group.

    Repeat steps 8.a – 8.e for each on the servers in the cluster.

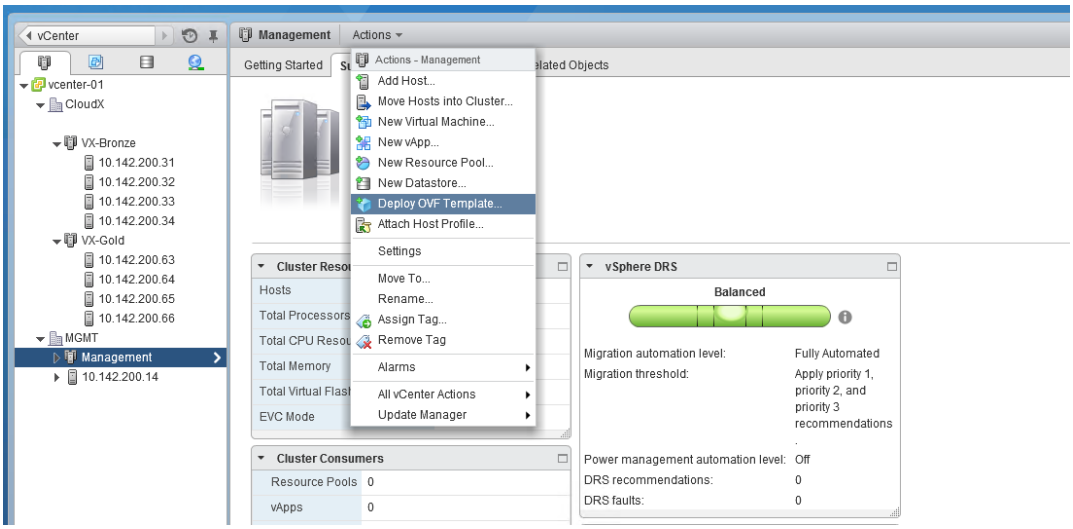
    Once completing these steps, the window below will display your results.



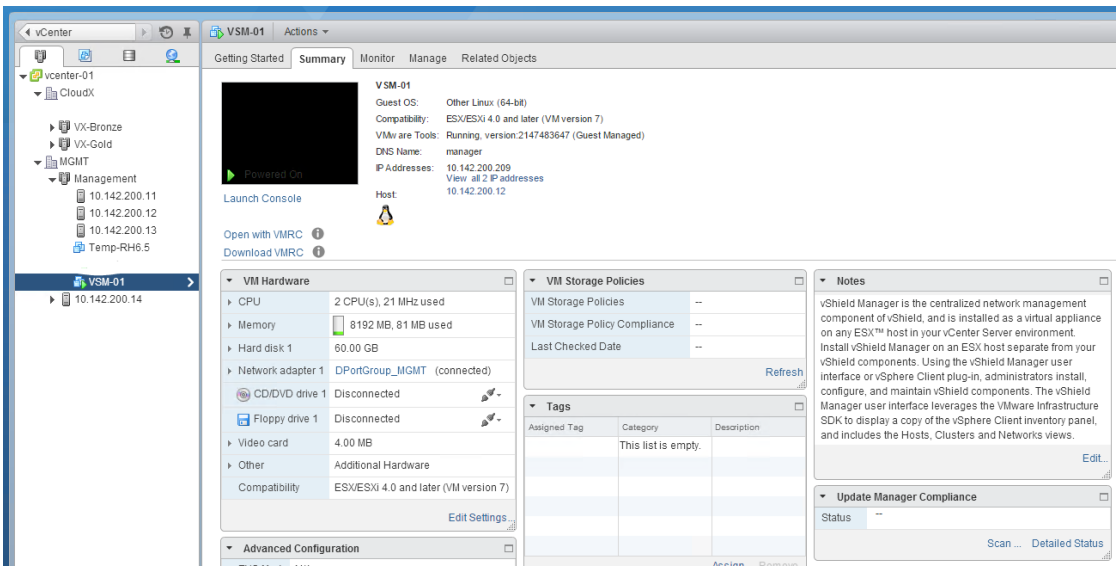
Mellanox Technologies Confidential

4 vShield Manager Appliance Installation

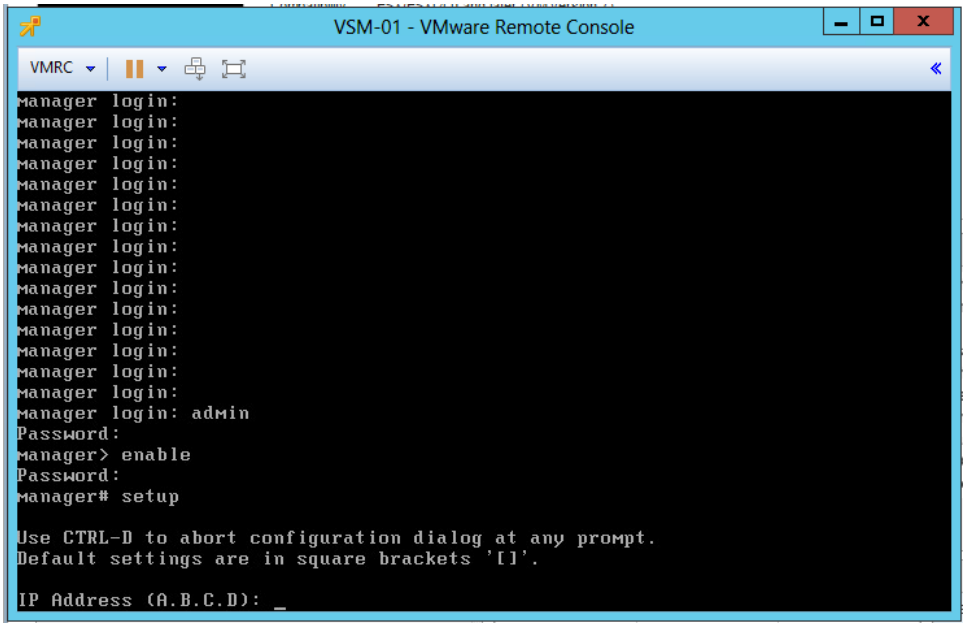
1. Download the vShield Manager Appliance from the VMware website.
2. Install the vShield Manager Appliance on the Management cluster, from the vShield Web Client.
   (Actions → Deploy OVF Template)



3. Complete the remaining steps in the wizard to install the vShield Manager Appliance to a location of your choosing. After finishing the setup wizard, a new vShield Manager VM should appear in your Hosts and Clusters inventory window.
4. Power on the Virtual Machine.



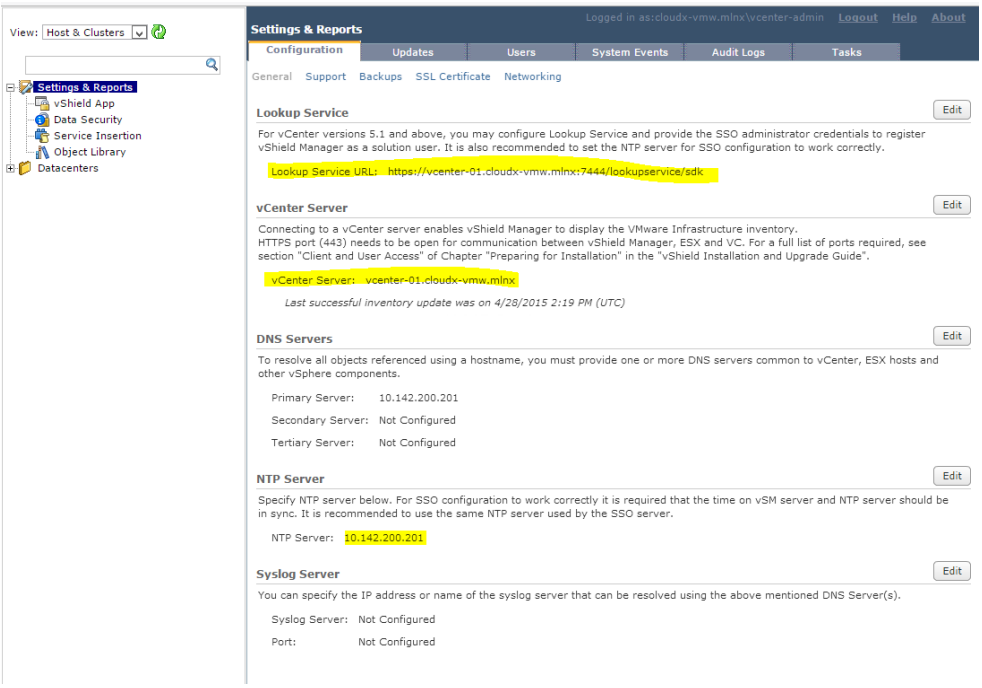
5. Click Open with VMRC.
6. Log in using **admin** credentials.
7. Type **enable** in the Command Line prompt and re-enter **admin** password.
8. Type **setup** to configure the vShield Manager VM networking settings in **Management and Storage** network.
   In this document, vShield Manager VM is referred as vsm-01.cloudx-vmw.mlnx.



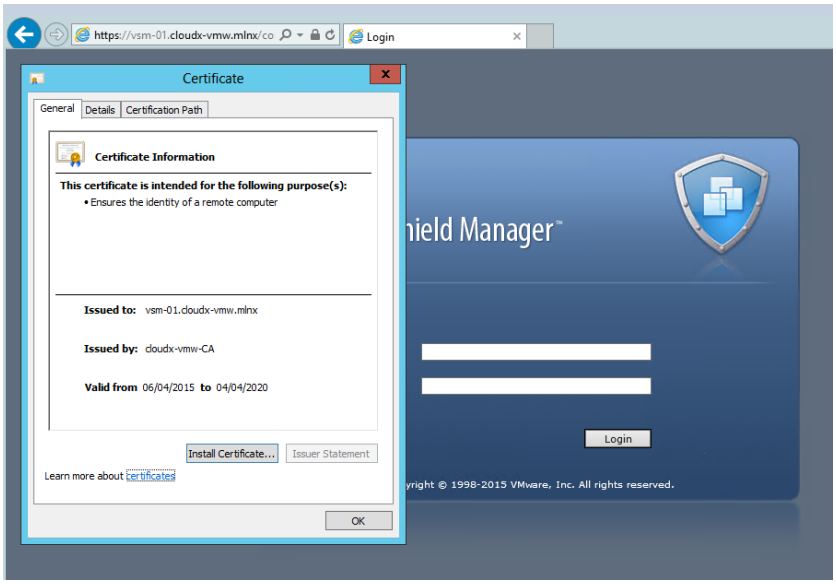
9. Open a web browser and navigate to: https://vsm-01.cloudx-vmw.mlnx

10. Log in using **admin** credentials.
11. Link the vShield Manager to your vCenter server.
    a. Navigate to Settings & Reports on the left pane. Under the Configuration tab on the right pane.
    b. Edit the information listed under vCenter Server to match your vCenter Server's network configuration.

Mellanox Technologies Confidential



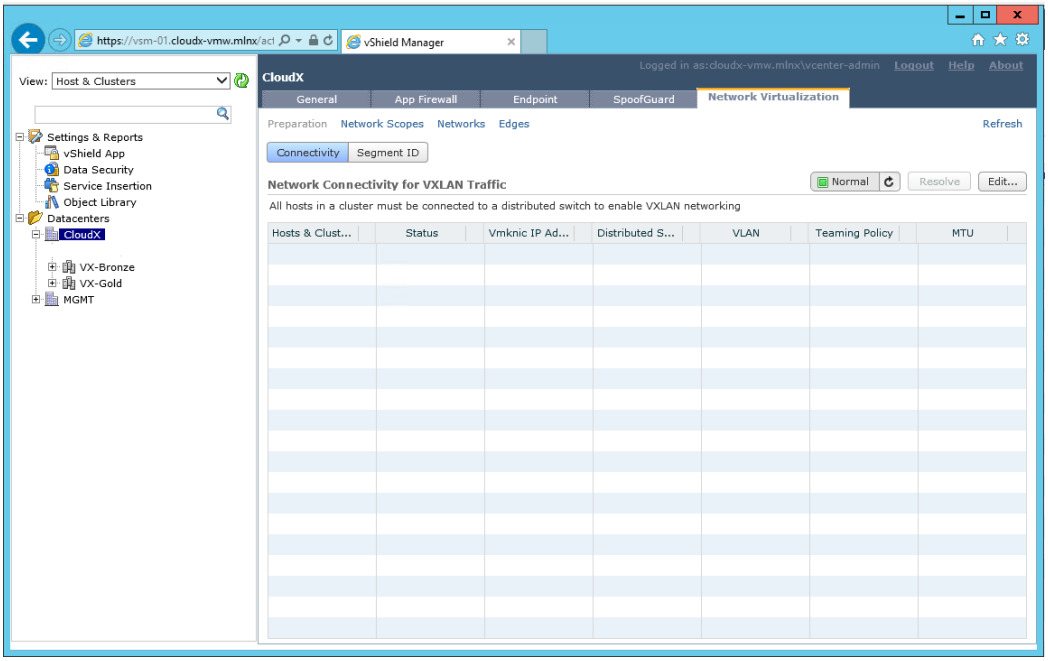
12. Add an SSL Certificate to identify the vShield Manager Web Service to disable the Certificate error warning.
    a. Click on the SSL Certificate under the Configuration tab on the right pane.
    b. Generate a Certificate Signing Request, download it, sign in a local CA and import the signed certificate.
13. Reboot the vShield Manager Appliance.
14. Re-open a web browser and navigate to: https://vsm-01.cloudx-vmw.mlnx
15. Verify the vShield Manager Web Service certificate is valid.



NOTE: For more information please refer to: http://www.vmware.com/pdf/vshield_55_install.pdf

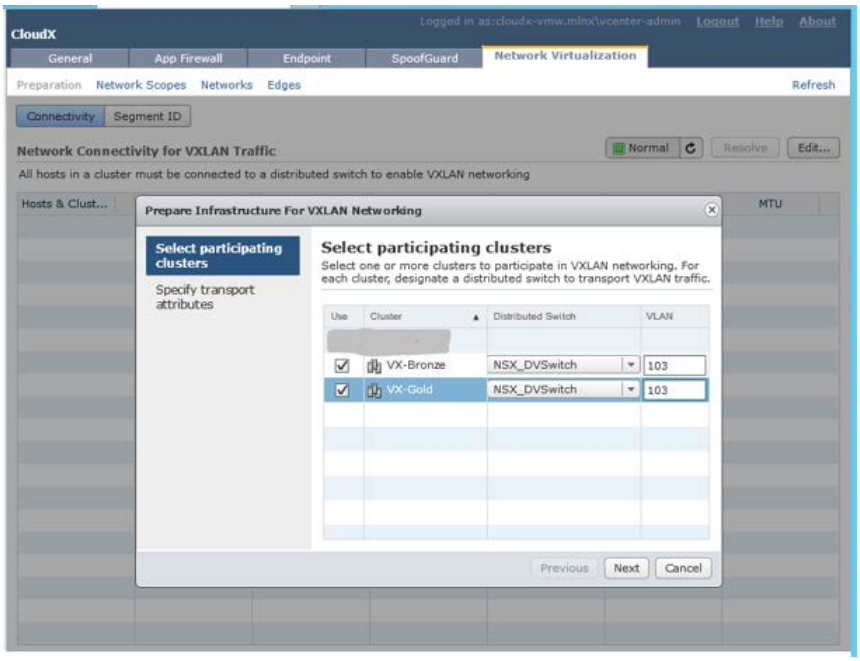
5 VXLAN Configuration

1. Log into the vShield Manager (VSM) via a web browser.
2. Select your CloudX datacenter in the left pane.
3. Navigate to the Network Virtualization tab on the right pane.



4. Press the **Preparation** link and add the Clusters you want to participate in VXLAN networking, and then press **Edit…**.

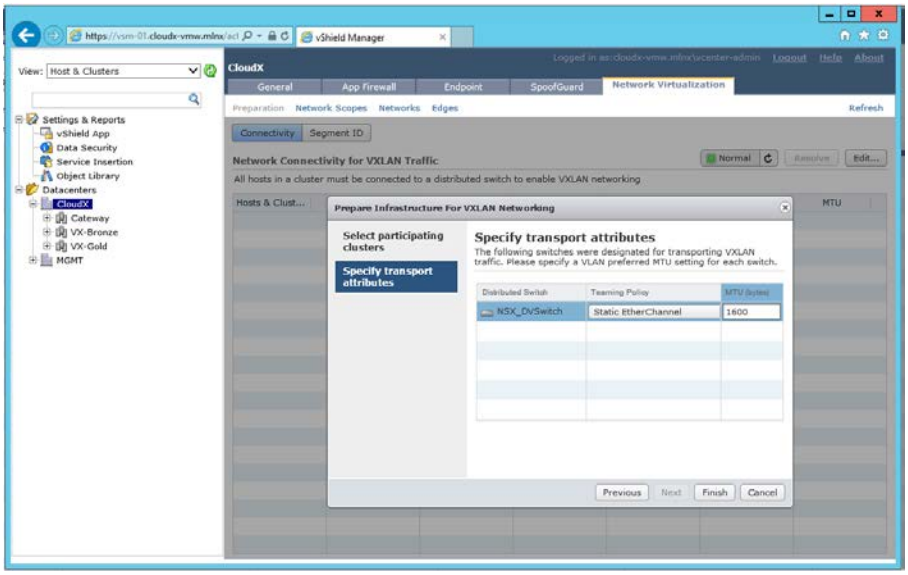
   a. Choose a distributed switch and set the VLAN ID.



   Click **Next**.

Mellanox Technologies Confidential

b. Set the desired teaming policy and MTU to 1600.
In this setup there is no teaming and servers connected through single 40Gbps link.

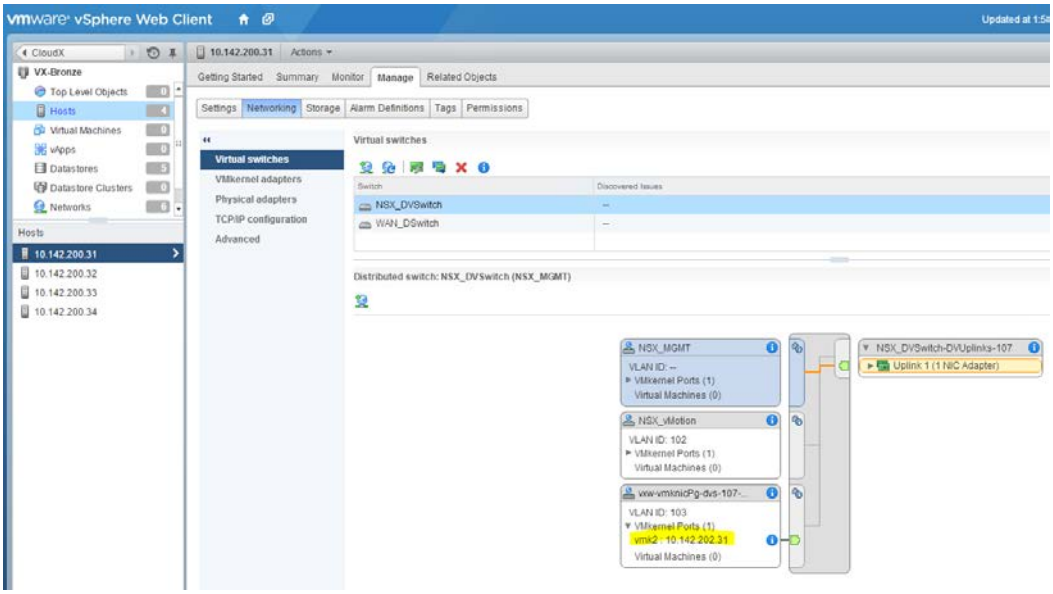


5. Click **Finish** to begin preparing your hosts for VXLAN networking.

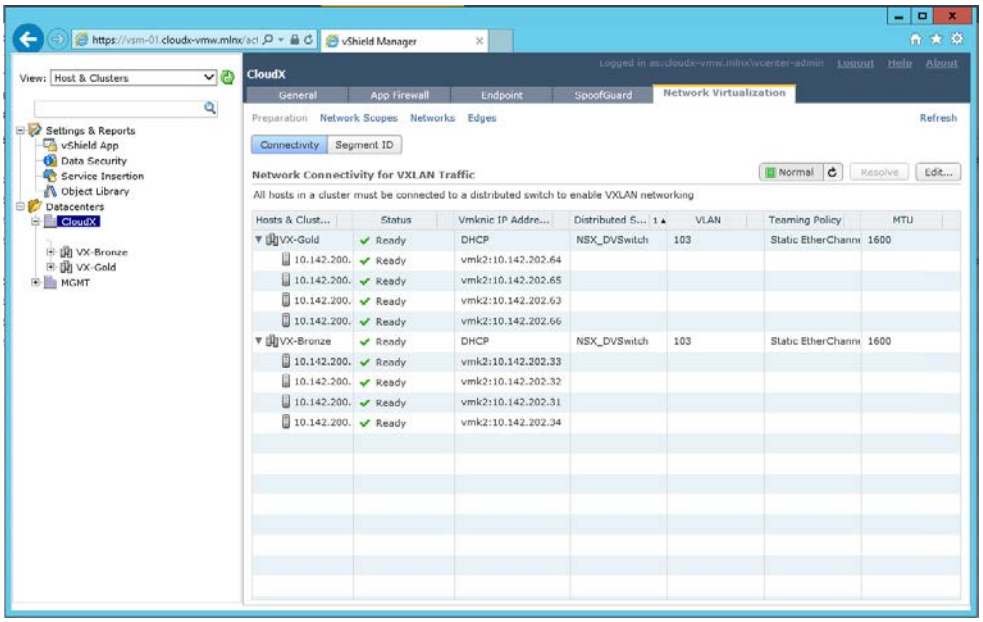
6 Virtual Tunnel Endpoint IP Configuration

1. The IP configuration is required for Virtual Tunnel Endpoint (VTEP) VMkernel interface on each host in CloudX datacenter. The configuration may be either provided by the DHCP server or configured manually.
   In this setup, there is no DHCP infrastructure available so VTEP VMkernel interfaces were configured manually.

   The figure below shows the final configuration for one VTEP VMkernel interface.

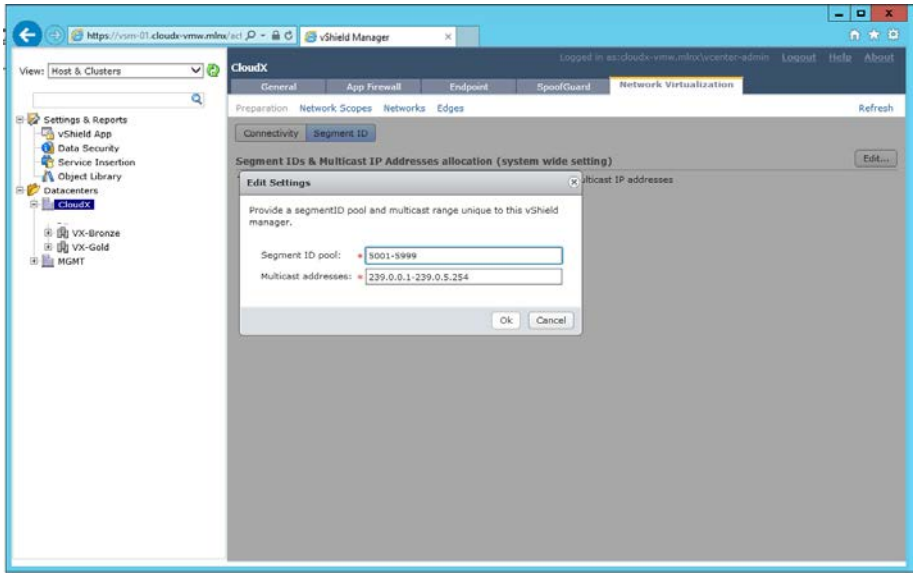


2. After setting up VTEP VMkernel interfaces on all hosts, refresh the Preparation page to see that all Clusters are VXLAN ready.



3. Set up the Segment ID Pool and Multicast Group Address Range configuration.

Click the **Segment ID** button and then press **Edit…**.



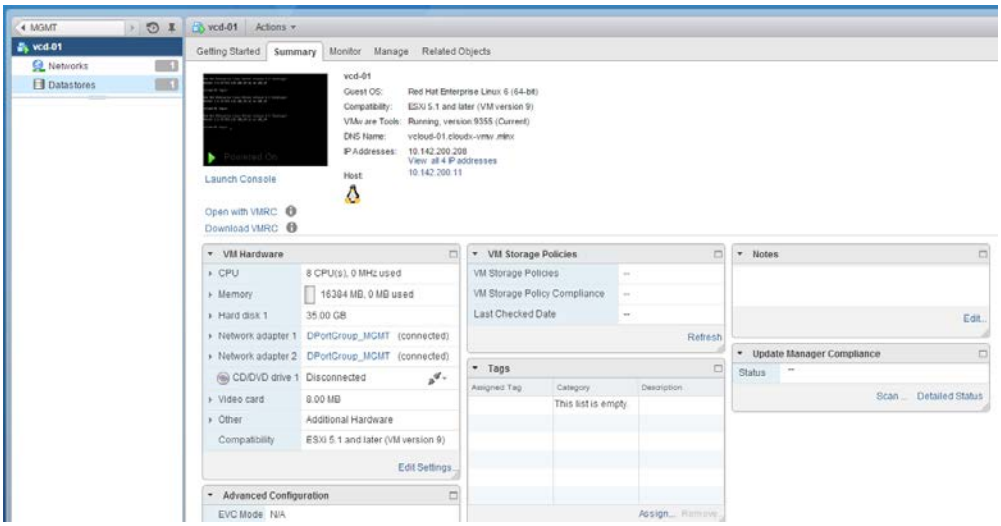
Click **OK**.

7 vCloud Director Deployment

1. Deploy vCloud Director.

   For information on how to do so, please refer to:
   http://pubs.vmware.com/vcd-55/topic/com.vmware.ICbase/PDF/vcd_55_install.pdf
   http://pubs.vmware.com/vcd-55/topic/com.vmware.ICbase/PDF/vcd_55_admin_guide.pdf

   The screenshot below displays the complete VM with the deployed vCloud director. In this setup, vCloud Director Hostname is **vcloud-01.cloudx-vmw.mlnx**.



2. After installing vCloud Director, open up a web browser and navigate to:
   https://vcloud-01.cloudx-vmw.mlnx/
   and complete the remaining steps in the wizard.